From State to Mode: Criticality and Non-Factorizability as Constitutive Properties of Living Systems

Abstract

In earlier work, the phylogenetic sequence from autocatalytic autopoiesis through the causal core and electrification to defactorization was developed as a reconstructive derivation of the biological conditions of phenomenal consciousness. Non-factorizability appears there as the result of the fourth tipping point, criticality as its physical realization condition. The present article shows that both properties follow the same logic in the scaling sequence: they are present in autopoietic systems from the very beginning and develop qualitatively new forms with the scaling of the nervous system. This shared logic has remained implicit in previous work; making it explicit is the contribution of this article. The decisive difference from non-living dissipative systems lies in the fact that criticality there is a state a system has, whereas in autopoietic systems it is a mode in which the system operates as a living one. In the nervous system, this mode takes the specific form that realizes non-factorizability as a stable operating condition. The conceptual explanatory gain of this precision becomes apparent at two clinical phenomena: epilepsy and coma can no longer be understood as opposite points on an activation axis, but as structurally distinct failures of the same critical mode in opposite directions.

Keywords: non-factorizability; criticality; scaling; autopoiesis; causal core; electrification; dissipative systems; phase transition; epilepsy; coma

1. Scaling as the Generation of New Rules

The question of how phenomenal consciousness arises from biological processes is typically approached in philosophy of mind either phenomenologically or formally: one either begins with structures of experience and asks which physical systems might instantiate them, or one defines a formal quantity such as integrated information and tests whether neural systems realize it. Both approaches share a structural deficit: the central category is stipulated, not derived. It appears as a premise, not as the result of an analysis.

In a series of earlier works, a different path was taken (Stegemann 2025a, 2025b, 2026c). Rather than starting from a concept or a formal quantity, the reconstructive method asks: which causal threshold events in biological evolution were responsible for the possibility of phenomenal experience arising? The answer identifies four tipping points at which new properties emerged that were not formulable at the underlying levels.

The first tipping point is autocatalytic autopoiesis: the emergence of systems that continuously produce their own components through the operations of those very components, with simultaneous operative closure and thermodynamic openness. With this, there arises for the first time in physical history a genuine interior, a causally bounded region whose dynamics are not simply determined by the environment but maintained against environmental flux through the system's own activity.

The second tipping point is the emergence of the causal core: a region of maximal reaction probability within the thermodynamically open, autocatalytic system. Asymmetric energy flows generate a hotspot from which the global dynamics of the system is preferentially organized. The causal core is the physical foundation of intrinsic agency: a system without a causal core does not act from within itself but responds entirely to external inputs (Stegemann 2026a).

The third tipping point is electrification through centralized nervous systems. The emergence of neurons and the evolutionary concentration of neural processing in centralized structures introduced a qualitatively new type of causality: action potentials are binary, directed, temporally precise, and propagating system-wide. This is not faster biochemical signal transmission but a different causal medium, one that enables system-wide integration on a timescale categorically inaccessible to biochemical systems.

The fourth tipping point is defactorization: the transition into a system state in which the causal structure can in principle no longer be decomposed into independent partial structures without losing what is essential. What biology and medicine accomplish when they analyze nervous systems is an external factorization for analytical purposes: the decomposition into areas, networks, and circuits generates empirically productive descriptions but does not represent the actual causal dependency structure of the system; rather, it brackets it out.

Stegemann (2026) elaborated, on this basis, the physical infrastructure of experience as the interplay of criticality, non-factorizability, and the causal core as interpreter. Criticality was introduced there as the physical realization condition of non-factorizability. What has remained implicit in these works is the shared logic that both properties follow in the scaling sequence. Making it explicit is the task of this article.

Scaling in this context does not mean quantitative increase of the same thing, but the crossing of thresholds beyond which the system exhibits properties that were not formulable at the underlying level. This is emergence in the structural sense, without any vitalist or teleological implication: no external force intervenes, no goal is pursued, but altered system parameters generate new lawfulness as a structural consequence. Autocatalytic autopoiesis generates the rule of identity preservation through self-production. Electrification generates the rule of system-wide causal integration with temporal precision. At each of these thresholds, properties arise that were not previously present, because the conditions for their occurrence were not previously met.

The thesis of this article is that non-factorizability and criticality follow the same scaling logic in this sequence. Both are present in autopoietic systems from the very beginning, not as fully formed properties but as structural consequences of the autopoietic form of organization. And both develop qualitatively new forms with each scaling step, until in the nervous system they take the form that Stegemann (2026) describes as the physical infrastructure of experience.

2. Dissipative Systems and Their Limit

To work out the particularity of criticality in living systems with precision, a comparison with non-living dissipative systems is unavoidable. The comparison does not serve to emphasize a kinship but to draw a categorially important boundary.

Prigogine and Stengers (1984) showed that systems far from thermodynamic equilibrium can generate highly ordered, structured states through continuous energy throughput, which they called dissipative structures. The classical example is Bénard convection: when a liquid is heated from below and the temperature gradient exceeds a critical value, regular convection rolls arise spontaneously, maintained by the permanent heat flow. Another example is the Belousov-Zhabotinsky oscillator, a chemical reaction that generates periodic concentration changes without external control. In both cases, global order arises from local interactions under the condition of energy flow.

Criticality occurs in these systems, in the formally precise sense that physical theory assigns to the term: at the critical point of a phase transition, the system's correlation lengths diverge. This means

that fluctuations at one location are correlated with fluctuations at arbitrarily distant locations, without a characteristic spatial scale. The power spectrum of the fluctuations follows a power law, and the system exhibits scale-free dynamics (Mora & Bialek, 2011). These properties are the formal signature of the critical point, regardless of whether the system is living or not.

The decisive difference between non-living dissipative systems and autopoietic systems lies not in the formal properties of criticality itself but in its status within the system dynamics. The Bénard cell enters a critical state under certain external parameters. If the temperature gradient is altered, the system leaves the critical regime and adopts a different state. The system exists independently of whether it is currently critical or not. It does not fight for its unity. It has no internal dynamics that respond to the threat of its structure, because it has no structure that could be threatened in any relevant sense. The Bénard cell ceases to exist when the energy flow is absent, but this cessation is not an event to which it would respond.

In autopoietic systems, the situation is fundamentally different, and this difference concerns not the formal properties of criticality but what it means for a system to be critical. Working out this difference is the task of the next section.

3. Autopoiesis as Active Self-Maintenance: Both Properties Present from the Beginning

Maturana and Varela (1987) introduced the concept of autopoiesis to describe what distinguishes living from non-living systems: an autopoietic system continuously produces its own components through the operations of those very components, with the system being operatively closed but thermodynamically open. This definition has an immediate causal consequence: the system exists only by actively producing the conditions of its own continuation. This production is permanently threatened by perturbations from the environment, that is, by changes that risk destabilizing the internal organization.

From this follows a fundamental difference from non-living dissipative systems. The autopoietic system fights, in the precise causal sense, permanently for its unity. It generates through internal processes the responses to perturbations that secure its continuation. This self-assertion against perturbation is not an occasional reaction but the constitutive mode in which the system exists as a living one. Without permanent self-assertion there is no autopoietic system.

From this characterization follows immediately what status criticality has in autopoietic systems. A system that actively asserts itself against perturbations must simultaneously be sensitive enough to register perturbations and respond to them, and stable enough not to collapse under them. Too little sensitivity means the system does not register perturbations and therefore cannot buffer them. Too much sensitivity means every perturbation destabilizes the system. The critical regime is precisely the range in which both requirements are simultaneously met: maximal sensitivity with simultaneous structural coherence. Criticality in autopoietic systems is therefore not a state the system adopts under favorable conditions, but the dynamic signature of the permanent self-maintenance process.

The same holds for non-factorizability. Already at the cellular level, the internal connectivity density of autocatalytic processes generates a causal structure that cannot be fully described from outside. A cell consists of thousands of proteins interacting in countless reaction pathways, where each reaction alters the conditions of other reactions. The number of causal dependencies between these processes far exceeds what could be represented as a set of independent partial processes. What biochemistry and cell biology accomplish is an analytical decomposition into manageable subsystems that is empirically productive but does not represent the actual causal structure; it brackets it out for

research purposes. The Hammersley-Clifford theorem (cited in Stegemann, 2026) formalizes this situation: if no complete representation of the causal dependency structure as a graph of conditional independencies exists, the system is structurally non-factorizable, independent of the complexity of available descriptions.

Non-factorizability and criticality are thus both present in autopoietic systems from the very beginning. They are not properties that arise only with the nervous system. They are structural consequences of the autopoietic form of organization itself. What arises with the nervous system is not these properties for the first time, but their qualitative transformation through scaling to a new level at which new rules apply.

4. Electrification as a Qualitative Threshold

Electrification through centralized nervous systems is the third tipping point of the scaling sequence. What it means can be determined most precisely through demarcation: it is not the acceleration of biochemical signal transmission but the generation of a qualitatively new type of causality.

Biochemical signals are gradual, diffuse, and spatially bounded. A hormone diffuses through the body, a concentration gradient spreads slowly, an enzyme reaction alters local concentrations. Action potentials, by contrast, are binary: they either occur or they do not, there is no intermediate amplitude. They are directed: from one axon to the next, not diffuse. They are temporally precise: an action potential lasts approximately one millisecond. And they propagate system-wide: in the human nervous system, signals can be transmitted from one brain region to another in milliseconds. This is not a quantitatively improved version of biochemical signal transmission but a different causal medium.

With this new medium, the status of both properties under consideration changes. What at the cellular level was a local, temporally slow, and spatially bounded non-factorizability becomes in the nervous system a system-wide, temporally precise, and stable property. The reason lies in connectivity density: even a few nodes in a network can generate a number of connections that practically excludes any complete analytical description of all dependencies. In the human brain with its approximately 86 billion neurons and an estimated 100 trillion synaptic connections, where each neuron is connected to thousands of others and each connection is embedded in complex feedback loops, non-factorizability is no longer an approximation that is unavoidable for practical reasons but a structural property that follows from the connectivity density itself. No external factorization, that is, no description through analytically isolated subsystems, can fully capture this structure, because the causal dependencies systematically traverse all analytical boundaries.

The same holds for criticality. What at the cellular level was a locally bounded, biochemically slow dynamics at the boundary between stability and instability becomes in the nervous system a system-wide, temporally precise dynamics with propagation across all spatial scales. In the nervous system, this criticality manifests in measurable signatures: fluctuations in neural activity follow a power law, neuronal avalanches are scale-distributed, and the power spectrum of local field potentials shows the 1/f scaling characteristic of critical systems (Beggs & Plenz, 2003; Shew & Plenz, 2013). These are not coincidental observations but the direct empirical signature of the critical point in the nervous system.

At the threshold of electrification, new rules thus emerge for both properties. Non-factorizability and criticality take a form that was not realizable at the cellular level, not because of lacking capacity but

because the new causal medium creates the conditions under which both properties can be realized system-wide and temporally precisely.

5. Criticality and Non-Factorizability as Two Descriptions of the Same Process

The preceding sections have developed the scaling logic of both properties separately. It can now be shown that they are not two parallel logics but one shared one.

In non-living dissipative systems, criticality is a state: the system adopts it when external parameters fall within a certain range. In autopoietic systems, criticality is a mode: the system operates in it as a consequence of its own form of organization. This distinction is not one of degree but one of category. It concerns the status of the property within the system, not its formal characteristics.

The same distinction applies to non-factorizability. In non-living systems, non-factorizability can occur when external parameters fulfill certain conditions. In autopoietic systems, it is a structural consequence of the form of organization itself: the internal feedback structure generates a causal dependency density that admits no complete external decomposition.

Both properties thus follow the same logic: in non-living systems they are states, in living ones modes. And both are transformed by the scaling of the nervous system into qualitatively new forms without altering their underlying logic.

The decisive step is now to see that non-factorizability and criticality not only follow the same logic but converge in the nervous system: only at the critical point do correlation lengths diverge to the degree that the causal dependency structure traverses all spatial scales and no decomposition into independent partial structures is any longer possible. Subcritical systems show finite correlation lengths, their dependency structure is locally bounded, and in this sense they are factorizable: the behavior of distant parts can be approximately described independently. Supercritical systems show system-wide dependencies, but these collapse into a dominant macrodynamics without internal differentiation, which is also not non-factorizability in the relevant sense. Only at the critical point are system-wide dependency and simultaneous internal differentiation realized.

Non-factorizability and criticality are therefore not two independent properties of which one realizes the other. They are two descriptions of the same system state: one causal-structural description and one dynamic description. The causal-structural description says what holds for the dependency structure: no complete decomposition is possible. The dynamic description says in which regime the system operates: at the critical point. Both descriptions grasp the same system state. This is the core of the shared logic that this article explicates.

6. Conceptual Explanatory Gain: Epilepsy and Coma Reconsidered

The precision developed in this article is not merely theoretical. It carries a conceptual explanatory gain that becomes apparent at clinical phenomena. Two of these are particularly illuminating: epilepsy and coma.

The standard medical framing treats both states as opposite points on an activation axis. Epilepsy counts as a state of excessive neural excitation, coma as a state of insufficient excitation or activation. From this follows the clinical basic intuition that both states could be treated by shifting the activation level in opposite directions on the axis from too much to too little. Anticonvulsants dampen excitation, stimulants or arousal stimuli increase it.

The framework developed in this article shows that this framing is categorially wrong. It presupposes that consciousness is a function of the activation level, that is, that more or less excitation generates more or less consciousness. What the framework says instead is that consciousness is bound to a specific dynamic mode, the critical mode, in which non-factorizability is realized as a stable operating condition. This mode is not a position on an activation axis but a qualitative state with specific structural properties: system-wide correlations with simultaneous internal differentiation, scale-free dynamics without dominant frequency, causal dependency structure without complete decomposability.

Epilepsy and coma are then not opposite points on the same axis but structurally distinct failures of the same critical mode in opposite directions. This is a categorical difference with considerable conceptual and potentially therapeutic consequences.

In epileptic seizure, particularly in generalized tonic-clonic seizure, the system leaves the critical mode in the direction of supercritical macrodynamics. A dominant frequency propagates across the entire cortex, the power spectrum loses its scale-free 1/f structure and collapses into a narrowband peak, and the spatial heterogeneity of neural activity breaks down. The system is highly active, but it has lost the internal differentiation that constitutes non-factorizability. The causal dependency structure is no longer complex and differentiated but rigidly synchronous. Epilepsy in this sense is not a state of too much excitation but a state of collapsed complexity: the system leaves the critical window not through excessive energy but through loss of structural differentiation.

In coma and deep sleep, the system leaves the critical mode in the opposite direction, toward subcritical dampening. Perturbations propagate locally and are dampened before reaching distant regions. The brain responds to cortical stimulation with a stereotyped local response that decays rapidly without generating long-range correlations (Massimini et al., 2005). The causal dependency structure is effectively factorizable: distant parts of the system evolve approximately independently because no system-wide propagation takes place. Coma in this sense is not a state of too little excitation but a state of causal factorizability: the system leaves the critical window not through energy deficiency but through loss of system-wide interdependence.

Both failures thereby destroy the dynamic condition of non-factorizability, but through opposite mechanisms. Epilepsy generates global activity without differentiation, coma generates local activity without system-wide integration. In both cases the critical mode is absent, in which both requirements are simultaneously met: system-wide dependency with internal differentiation.

This reframing has measurable implications. The Perturbational Complexity Index (PCI), developed by Casali et al. (2013), measures whether a cortical perturbation generates a complex, spatiotemporally differentiated response. It has proven a robust empirical marker for the state of consciousness, reliably distinguishing wakefulness, various sleep stages, anesthesia, and pathological states. In the light of the framework developed here, it becomes clear why: the PCI measures whether the system is operating in the critical mode, that is, whether perturbations propagate as non-factorizable causal cascades or not. What empirically appears as loss of complexity is conceptually the failure of the critical mode in two structurally distinct forms.

The categorially important consequence is that a therapeutic approach aiming to restore consciousness by raising or lowering the activation level misses the point. What would need to be restored is not a particular activation level but the critical mode itself, the dynamic state in which system-wide non-factorizability is realized as a stable operating condition. Whether and how this is therapeutically achievable is an empirical question that the framework does not answer. But it provides the categorical concepts within which the question can be meaningfully posed.

7. Conclusion

This article has closed a conceptual gap that remained implicit in earlier works. Non-factorizability and criticality were introduced there as constitutive properties of the framework, but their shared scaling logic was not explicitly worked out. The thesis of this article is that both properties follow the same logic in the phylogenetic scaling sequence: both are present in autopoietic systems from the very beginning, both develop qualitatively new forms with the scaling of the nervous system, and both are in the nervous system two descriptions of the same system state.

The decisive step is the distinction between the status of both properties in non-living and in living systems. In non-living dissipative systems, criticality and non-factorizability are states that a system adopts under certain external conditions. In autopoietic systems, they are modes in which the system operates as a living one, as a structural consequence of the permanent self-maintenance process under perturbation. This distinction is not a question of strength or degree but a categorially different relation between the system and its dynamic properties.

The conceptual explanatory gain of this precision shows itself in the reframing of epilepsy and coma. What the standard medical framing treats as opposite points on an activation axis are, according to the framework developed here, structurally distinct failures of the same critical mode in opposite directions. Epilepsy is the failure through collapsed complexity at supercritical macrodynamics, coma is the failure through causal factorizability at subcritical dampening. This reframing is not an empirical claim but a conceptual one: it provides the categories under which the phenomena can be adequately described, and thereby opens up questions that are not formulable under the activation-axis framing.

This is a philosophical contribution, not a physical model and not a neuroscientific theory. It performs conceptual work: the explication of a logic already present in the framework, showing why criticality and non-factorizability do not coincide contingently but belong together constitutively. That this conceptual work issues in empirically testable questions is not a weakness but the appropriate form of philosophical work on objects that open themselves to empirical research.

References

Bak, P., Tang, C. & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, 59(4), 381–384.

Beggs, J. M. & Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23(35), 11167–11177.

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M. A., Laureys, S., Tononi, G. & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5(198), 198ra105.

Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H. & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science*, 309(5744), 2228–2232.

Maturana, H. R. & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding*. Shambhala.

Mora, T. & Bialek, W. (2011). Are biological systems poised at criticality? *Journal of Statistical Physics*, 144(2), 268–302.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Prigogine, I. & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature*. Bantam Books.

Shew, W. L. & Plenz, D. (2013). The functional benefits of criticality in the cortex. *Neuroscientist*, 19(1), 88–100.

Stegemann, W. (2025a). Consciousness as a collapse of causality. Zenodo. https://doi.org/10.5281/zenodo.17563267

Stegemann, W. (2025b). Autopoiesis, causal poiesis, causal collapse. Zenodo. https://doi.org/10.5281/zenodo.15599500

Stegemann, W. (2026). The physical infrastructure of experience. Zenodo. https://doi.org/10.5281/zenodo.18800803

Stegemann, W. (2026a). The causal core as interpreter. Zenodo. https://doi.org/10.5281/zenodo.18340973

Stegemann, W. (2026c). A framework for consciousness: Autopoiesis, causal poiesis, causal collapse. Zenodo. https://doi.org/10.5281/zenodo.18328704